4-6 October 2023, Universidad Carlos III de Madrid, Spain

Book of Abstracts

Steering Committee

Vincent Blondel (University of Louvain)
Leo Ferres (IDSUDD, ISI Foundation & Telefónica)
Marco Fiore (IMDEA Networks Institute & Net AI)
Vanessa Frias-Martinez (University of Maryland)
Renaud Lambiotte (University of Oxford)
Esteban Moro (MIT and Universidad Carlos III de Madrid)
Alex 'Sandy' Pentland (MIT)

Main conference

Talks

Mind the Gap: Studying the Impact of Covid-19 on Cellular Users Presence in Inland Areas

Andrea Pimpinella[1], Carmelo Ignaccolo[2], Cristina Boniotti[3]
[1] *Department of Electronics, Information and Bioengineering (DEIB), Politecnico di Milano*
[2] *Department of Urban Studies and Planning (DUSP), Massachusetts Institute of Technology*
[3] *Department of Architecture, Built Environment and Construction Engineering (DABC), Politecnico di Milano*
email: {name.surname}@polimi.it[1,3], carmeloi@mit.edu[2]

I. Introduction

It is widely known that there is a close relationship between the habits of people and the characteristics of the environment where such people live. Also, due to the ubiquity of mobile radio access (mobile subscriptions are expected to overcome 9 billions in 2028 [1]), the study of cellular users communication activity can reveal important insights about social, topological, and technological phenomena at large scales [2].

The goal of grouping mobile radio access sites (e.g., eNodeBs) according to the spatial and temporal characteristics of their network activity is typically pursued through the design of *clustering* algorithms. A common choice is to group network sites according to the dynamics of the served traffic [3]: this option offers several advantages to service providers, who put effort to discover regularities of traffic loads and target pro-activity in network resources management.

Recently, the huge potential of mobile data has been exploited to examine the impact of Covid-19 pandemic on people's mobility and social life[1]. In this context, many studies in literature observe a shift in the presence of people from dense urban cities to inner areas, apparently inverting the trend of depopulation of rural areas that has been observed since decades in many countries worldwide [4], [5].

Using a real-world cellular network dataset, this study focuses on the effects of Covid-19 pandemic on cellular users presence in the Italian region of Valtellina, which consists of small, rural villages progressively shrinking due to a long-term depopulation phenomenon. By clustering network sites based on the changes (i.e., gaps) in the relative share of connected users after the pandemic, we observe that: i) users' behavior in Valtellina has changed and ii) such change is spatially heterogeneous, indicating a modification in the attractiveness of urban settlements according to post-pandemic needs.

II. Data and Methodology

This work leverages a dataset coming from the LTE network of a popular European mobile operator, containing radio access network measurements collected at 61 eNodeBs in the region of Valtellina, in the form of hourly sampled time series. We consider here the number of users that are Radio Resource Control (RRC) connected to each cell site as a proxy of people presence in the valley. We focus on two 1-month periods, T_1={20/01/2020, 16/02/2020} (i.e., before pandemic breakout in Italy[2]) and T_2={24/01/2022, 20/02/2022} (i.e., when national government removed most of Covid-19 restrictions).

Operatively, we rely on k-means clustering to group together eNodeBs based on the Euclidean distance among their *gap* signals. In detail, the gap signal $\mathbf{g}_i$ of the i-th eNodeB is generated as it follows:

1) Compute the Median Weekly Signature (MWS) of the number of connected users in T_1 and T_2 (computational details can be found here [3]), namely $\mathbf{m}_{(1,i)}$ and $\mathbf{m}_{(2,i)}$.
2) Normalize each hour of both signatures to the hourly sum of connected users, referring to them as $\mathbf{m}^n_{(1,i)}$ and $\mathbf{m}^n_{(2,i)}$.
3) Compute $\mathbf{g}_i$ as the difference between $\mathbf{m}^n_{(1,i)}$ and $\mathbf{m}^n_{(2,i)}$.

Repeating this process for each eNodeB in the network allows to perform k-means clustering with the gap signals $\mathbf{g}_i$ as input, setting $k = 4$ through a data-driven approach [3]. We underline that our approach takes into account both the spatial and temporal dimensions of the problem. On the one hand, we consider the share of the total number of cellular users connected in Valtellina taken every hour by each site through the normalization process as described above (step 2), i.e., we embed spatial domain information in the clustering objects. On the other hand, the algorithm clusters $\mathbf{g}_i$ according to its temporal dynamics, thus using the time domain information of the signals. Note also that each input is normalised to mean and variance in the time domain before being used, as our interest is to group eNodeBs regardless of the amplitude of the users' presence variation.

III. Experimental Results

We plot in Figure 1 the centroids of the recognised clusters, each one representing the gap signals of all eNodebs in the corresponding cluster. For each centroid, positive (negative) gaps represent the case where the variation of the number of users served from T_1 to T_2 at that hour is greater (smaller) than the centroid's average weekly gap, that is represented by the black horizontal dashed line.

Moving downwards in Figure 1 from top to bottom plot, we interpret the centroid profiles as it follows:

1) The red cluster groups 20 eNodeBs (33% of the total), and is the largest of our configuration. Its centroid has

[1]Examples can be found at these web sites: i) https://senseable.mit.edu/stockholm-19/, ii) https://wfh-inequalities.netlify.app

[2]First recognized Covid-19 case in Italy dates to 20/02/2020.

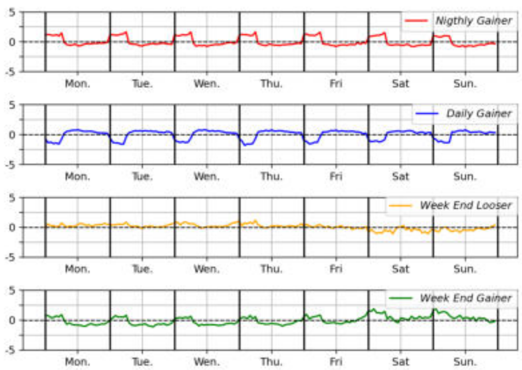

Fig. 1. Weekly profiles of centroids representative of the recognised clusters.

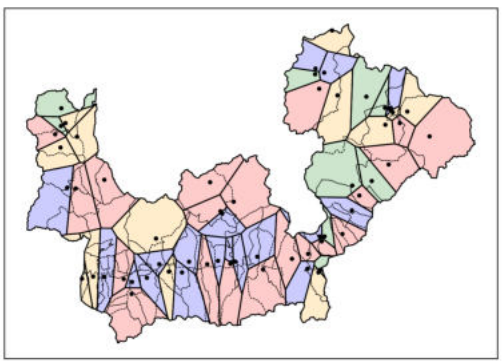

Fig. 2. Geographical map of Valtellina. Voronoi polygons are color-coded according to the cluster the corresponding eNodeB belongs to.

high gaps during the night (from 0:00 a.m. to 7:00 a.m.), while they are slightly below the average during the day, for both week days and week ends. We therefore label the eNodeBs of this cluster as *Nightly Gainers*.

2) The blue centroid repesents 19 eNodeBs (31%), which we name as *Daily Gainers*. In fact, we observe gaps greater than the weekly average from 9:00 a.m. to 11:00 p.m in week days (up to 8:00 p.m. in the week end), while they lay below the average during the night.
3) The orange centroid, representing 13 out of 61 eNodeBs (21%), has a profile that is stable around the average during week days. Differently, gaps turn negative during the week end: this means that the eNodeBs of this cluster have more likely lost users during the week end in T_2 as compared to T_1 than what observed during the week days. So, we label such eNodebs as *Week End Loosers*.
4) The green cluster contains only 9 out of 61 eNodeBs (14%), and is the smallest of the four clusters. As one can see, its centroid has a week day profile similar to the one of the red cluster but with a positive upward trend. In fact, gaps are positive during the week end: we thus name the eNodeBs in this cluster as *Week End Gainers*.

We summarise in Table III the percentage of eNodeBs of the network grouped in each cluster and the corresponding fraction

TABLE I
PER-CLUSTER DISTRIBUTION OF ENODEBS AND CUMULATIVE SHARE OF CONNECTED USERS DURING T_2.

Cluster Label	**Color**	**Cluster Members (%)**	**Served Users (T_2) (%)**
Nightly Gainer	Red	32.79	30.73
Daily Gainers	Blue	31.15	42.92
Week-End Gainer	Orange	21.31	14.24
Week-End Looser	Green	14.75	12.12

of the overall number of connected users that were served during T_2. To understand the spatial distribution of the clusters, we represent in Figure 2 a colour-coded map of Valtellina, where each color corresponds to the mentioned clusters. In the map, the black markers pinpoint the eNodeBs location, while the thicker black boundaries outline the radio coverage area of each eNodeB, which for simplicity is here assumed to coincide with the corresponding Voronoi polygon[3]. Also, thinner black lines outline the administrative boundaries of the municipalities in the valley. As one can see, most of the Daily Gainers (blue polygons) are located in the middle Valtellina. In fact, municipalities in this area have recently upgraded their digital infrastructures thanks to recent public investments[4]: we believe communications technology upgrades greatly benefit users enjoyement, especially due to the enabling of remote working options. Additionally, we recall that blueish eNodeBs gain users also during the week end: this can be the results of the many efforts local public entities devoted to foster tourism in these areas. Besides, lateral valleys mostly host Nightly Gainers (red polygons), probably due to the scarcity of digital infrastructures and services as well as job opportunities, which induce citizens to move outwards during daily hours and return home during the night. Finally, we mention that the municipality of Grosio (whose administrative boundaries mostly overlap the green Voronoi polygon located eastwards close to Swiss border) is a Week End Gainer: it became in 2020 a neutral delivery point for optical cables distribution to surrounding districts[5]. Future works will regard a more specific analysis on the relationship between these findings and the characteristics of the built environment in the valley.

REFERENCES

[1] "Ericsson mobility report," https://www.ericsson.com/en/reports-and-papers/mobility-report, 2022.

[2] R. Singh *et al.*, "Urban vibes and rural charms: Analysis of geographic diversity in mobile service usage at national scale," in *The W.W.W. Conf.*, 2019, pp. 1724–1734.

[3] A. Pimpinella *et al.*, "Forecasting busy-hour downlink traffic in cellular networks," in *ICC 2022 - IEEE Int. Conf. on Communications*, 2022, pp. 4336–4341.

[4] G. Lanza *et al.*, "Impacts of the covid-19 pandemic in inner areas. remote work and near-home tourism through mobile phone data in piacenza apennine," *TEMA*, vol. 2, pp. 73–89, 2022.

[5] E. Willberg *et al.*, "Escaping from cities during the covid-19 crisis: Using mobile phone data to trace mobility in finland," *ISPRS Int. Journal of Geo-Inf.*, vol. 10, no. 2, p. 103, 2021.

[6] O. E. Martínez-Durive *et al.*, "Voronoiboost: Data-driven probabilistic spatial mapping of mobile network metadata," in *2022 19th Annual IEEE Int. Conf. on Sensing, Communication, and Networking (SECON)*. IEEE, 2022, pp. 100–108.

[3] How to represent the coverage area of a radio access point without the help of ad-hoc coverage maps is still a topic of literature debate [6].

[4] Ultra-Broadband National Plan (2015) and National Recovery and Resiliency Plan (2021).

[5] https://bandaultralarga.italia.it/en/map/